Software Architecture Working Group   **March 01, 2010**

Minutes of February 18, 2010 Meeting

Agenda

1.   Updates and Announcements

2.   R5.1.1. – code freeze/release date

3.   Google scholar progress

4.   Review of annotations specification

5.   Rights-event/document linking capability

Announcements and Quick Updates

Ron reported that year-end funds were being considered for the following (in priority order): 1) AC for the Systems server room, 2) additional public PCs, and 3) replacement of the PCs in the SCC IHLs. The work we have done in R5.1 to improve performance is obvious from Google page loading data. Jeffery has reported on a graph that shows RUcore page loading times to be, on average, 1.8 seconds – performance that is better than 70% of the sites monitored. From the graph, it also shows that our page loading times dropped dramatically in late December – coincident with our release of R5.1. From the previous NJVid conference call, it was noted there is a requirement to upload multiple PDFs with a video. Kalaivani indicated that we cannot currently do this and it should be reflected as a general requirement for our content models, e.g. this capability will probably be needed for datasets. Kalaivani has informed Tom Nemeth and Sujay that additional PDFs can be added via dlr/EDIT. Regarding the issue of having multiple copies of the same object with the handle pointing to the most recent, Kalaivani will delete the extraneous objects. We will also plan to continue to run a script to check for this condition, although we believe this problem is no longer occurring (probably due to early use of WMS and people doing multiple ingests of the same object).

R5.1.1 Review

In a review of outstanding bugs to be fixed, it appears that Jeffery and Jie have completed their work and Chad can finish within a week or so. There are some 25 outstanding bugs for WMS and Yang would probably need several months to complete all of the bugs. As a result, Kalaivani agreed to prioritize the 25 WMS bugs into two clumps. So, for example, we might fix the high priority clump in a R5.1.1 release and then pick up the second clump in the next full release (R5.2). (Note, this schedule is probably in jeopardy given that Yang has to rebuild his development environment due to a workstation crash).

Google Scholar

Jeffery reported that progress has been made and he has been able to contact a real person at Google. This person has indicated that additional meta-tags are need to include a resource in Google Scholar. This approach looks fairly straightforward and we will plan to include it in R5.2. It looks like only a change in "showfed" at this point. Jeffery will continue to investigate an approach where we provide a sitemap to Google Scholar and also determine if Google Scholar can use our Handle urls. Our first focus will be on our ETD collection. If we can get

this working, we will then proceed to other collections (e.g. faculty deposits). It should be noted that some of our ETDs are being picked up by Google Scholar now, probably because they have been linked from other sites.

Annotation Specification

This was the third (and final) review of the annotation specification. Only the major points are summarized here. Chad will update the document and also note where major modifications to other parts of RUcore software will be needed in the R5.2 release.

- There will be no impact on WMS in R5.2. For example, all ingests and edits will be handled by the Annotation function.
- Showfed will need to be modified for the following: a) to handle many annotations for a resource and b) to sort the annotations for presentation – possibly by time code.
- Indexing. We agreed that annotations will not appear in the search results list. However all annotation content will be indexed and included as part of the resource object. We need to determine where this indexing flag appears – perhaps with a relationship registry database (TBD). Note that this capability will also be needed for compound data objects.
- Policies (xacml). The xacml will be inline in the annotation object and we will need the annotator's ID. This ID will need to be extracted from LDAP and inserted into policy. In addition, we will need to add the ID to the user XML file (the xacml specification will be reviewed in the next sw_arch meeting.)
- Shibboleth. As part of this discussion, we agreed to turn on Shibboleth in our next release. (Post-meeting note. We need more discussion of how Shibboleth would be used in R5.2).

Other Items

Sho mentioned that there is a one line change for configuring Fedora for UTF-8 (related to character translation problems). Jeffery will update the configuration document to reflect this change. There was discussion of how we can have more parallelism in our development process. For example, people who have finished a development for a specific release can either work on the backlog of bugs or alternatively start development work on the next release. To start on the next release, we will also need more parallelism in our requirements review process. (As a post-meeting editorial comment, we have difficulty getting through all of our requirement reviews in a timely fashion, in part because it takes a fair amount of meeting time in the software architecture group to deal with the technical complexities of a particular issue. One approach is to do more of the review outside of our primary meetings and just review the results/conclusions in software architecture. More discussion is needed.)

Agenda Items for the Next Meeting

- Release R5.1.1 and bug fix review
- Relationship services in R5.2
- Google scholar
- WMS/policy (xacml) integration
- Rights-event/document linking capability
- Faculty survey requirements

- Batch ingest for NJEDL – statistics and process
- Storage management on lefty64

rcj – 03/01/2010